Finite Fields and Their Applications **7,** 92–109 (2001)
doi.10.1006/ffta.2000.0302, available online at http://www.idealibrary.com on IDEAL®

Galois Groups of Generalized Iterates of Generic Vectorial Polynomials[1]

Shreeram S. Abhyankar

Mathematics Department, Purdue University, West Lafayette, Indiana 47907
E-mail: ram@cs.purdue.edu

and

Ganapathy S. Sundaram

Bell Labs, Lucent Technologies, New Jersey 07981
E-mail: ganeshs@bell-labs.com

Communicated by Daqing Wan

Received February 2, 2000; published online November 21, 2000

Let $q = p^u > 1$ be a power of a prime p, and let k_q be an overfield of GF(q). Let $m > 0$ be an integer, let J^* be a subset of $\{1, \ldots, m\}$, and let $E^*_{m,q}(Y) = Y^{q^m} + \sum_{j \in J^*} X_j Y^{q^{m-j}}$ where the X_j are indeterminates. Let $J^\ddagger$ be the set of all $m - v$ where v is either 0 or a divisor of m different from m. Let $s(T) = \sum_{0 \le i \le n} s_i T^i$ be an irreducible polynomial of degree $n > 0$ in T with coefficients s_i in GF(q). Let $E^{*[s]}_{m,q}(Y)$ be the generalized sth iterate of $E^*_{m,q}(Y)$; i.e., $E^{*[s]}_{m,q}(Y) = \sum_{0 \le i \le n} s_i E^{*[i]}_{m,q}(Y)$, where $E^{*[i]}_{m,q}(Y)$, is the ordinary ith iterate. We prove that if $J^\ddagger \subset J^*$, m is square-free, and $\mathrm{GCD}(m, n) = 1 = GCD(mnu, 2p)$, then $\mathrm{Gal}(E^{*[s]}_{m,q}, k_q(\{X_j : j \in J^*\}) = GL(m, q^n)$. The proof is based on CT (= the Classification Theorem of Finite Simple Groups) in its incarnation as CPT (= the Classification of Projectively Transitive Permutation Groups, i.e., subgroups of GL acting transitively on nonzero vectors). © 2000 Academic Press

1. INTRODUCTION

Throughout this paper let $q = p^u > 1$ be a power of a prime p, let $m > 0$ and $n > 0$ be integers, and let $GF(q) \subset k_q \subset K \subset \Omega$ be fields where Ω is an

[1]1991 Mathematical Subject Classification: 12F10, 14H30, 20D06, 20E22. Abhyankar's work was partly supported by NSF Grant DMS 97-32592 and NSA grant MDA 904-97-1-0010.

1071-5797/00 $35.00

algebraic closure of K. Also let $E = E(Y)$ be a monic separable vectorial q-polynomial of q-degree m in Y over K; i.e.,

$$E = E(Y) = Y^{q^m} + \sum_{i=1}^{m} X_i Y^{q^{m-i}} \quad \text{with} \quad X_i \in K \quad \text{and } X_m \neq 0, \tag{1.1}$$

where the elements $X_1, \ldots, X_m$ need not be algebraically independent over k_q. When we want to assume that, for a sequence of integers $J^* = (J_1^*, \ldots, J_{W^*}^*)$ with $1 \leq J_1^* < \ldots < J_{W^*}^* \leq m$, the elements $X_{J_1^*}, \ldots, X_{J_{W^*}^*}$ are algebraically independent over k_q and $K = k_q(X_{J_1^*}, \ldots, X_{J_{W^*}^*})$ with $X_i = 0$ for all $i \notin \{J_1^*, \ldots, J_{W^*}^*\}$, we may express this by saying that we are in the *generic* case of type J^*, and we may indicate it by writing $E_{m,q}^*$ for E and K^* for K. When J^* is the complete sequence $J^\sharp = (1, \ldots, m)$, we may say that we are in the *very generic* case to mean that we are in the generic case of type $J^\sharp$, and we may indicate this by writing $E_{m,q}^\sharp$ for E and $K^\sharp$ for K. When J^* is the singleton $J^\flat = (m)$ we may say that we are in the *binomial* case to mean that we are in the generic case of type $J^\flat$, and we may indicate this by writing $E_{m,q}^\flat$ for E and $K^\flat$ for K. When J^* is the pair $J^\dagger = (m-1, m)$ with $m > 1$ we may say that we are in the *trinomial* case to mean that we are in the generic case of type $J^\dagger$, and we may indicate this by writing $E_{m,q}^\dagger$ for E and $K^\dagger$ for K. Finally, when J^* is the sequence $J^\ddagger = (m - v_1, \ldots, m - v_W, m)$, where $m = v_0 > v_1 > \cdots > v_W = 1$ are all the positive divisors of m, we may say that we are in the *divisorial* case to mean that we are in the generic case of type $J^\ddagger$, and we may indicate this by writing $E_{m,q}^\ddagger$ for E and $K^\ddagger$ for K; note that if $m = 1$ then we are in the divisorial case with $W = 0$. As usual, for any other sequence of integers $J^{**} = (J_1^{**}, \ldots, J_{W^{**}}^{**})$ with $1 \leq J_1^{**} < \cdots < J_{W^{**}}^{**} \leq m$, we write $J^{**} \subset J^*$ or $J^* \supset J^{**}$ to mean that the sequence J^{**} is a subsequence of the sequence J^*. In particular, we always have $J^* \subset J^\sharp$. Likewise, if $J^\dagger \subset J^*$ then we must have $m > 1$. Moreover, if m is prime then $J^\dagger = J^\ddagger$. Similarly, if $m = 1$ then $J^\flat = J^\ddagger = J^\sharp$.

In the *general* (= not necessarily generic) case, let V be the set of all roots of E in Ω, and note that then V is an m-dimensional $\mathrm{GF}(q)$-vector-subspace of Ω. Moreover, since $\mathrm{GF}(q)$ is assumed to be a subfield of k_q and hence of K, every K-automorphism of the splitting field $K(V)$ of E over K induces a $\mathrm{GF}(q)$-linear transformation of V. Consequently $\mathrm{Gal}(E, K) < \mathrm{GL}(V)$; i.e., the Galois group of E over K may be regarded as a subgroup of $\mathrm{GL}(V)$ (see [Ab3]). We use $<$ to denote subgroup; some authors use $\leq$. If we do not assume $\mathrm{GF}(q) \subset k_q$ then we only get $\mathrm{Gal}(E, K) < \Gamma\mathrm{L}(V)$, where $\Gamma\mathrm{L}(V)$ is the group of all semilinear transformations of V (see [Ab6]). By fixing a basis of V we may identify $\mathrm{GL}(V)$ with $\mathrm{GL}(m, q)$, and $\Gamma\mathrm{L}(V)$ with $\Gamma\mathrm{L}(m, q)$.

In [Ab2, Ab3] it was shown that in the trinomial case we have

$$\mathrm{Gal}(E_{m,q}^\dagger, K^\dagger) = \mathrm{GL}(m, q). \tag{1.2}$$

For a somewhat simpler proof of (1.2) see [Ab4], and for applications of it see [Ab5]. Using specialization, where we put $X_i = 0$ for $i \notin J^\dagger$, from (1.2) we deduce that in the generic case of type J^* we have

$$J^\dagger \subset J^* \Rightarrow \mathrm{Gal}(E^*_{m,q}, K^*) = \mathrm{GL}(m,q) \tag{1.3}$$

and consequently in the divisorial case we have

$$\mathrm{Gal}(E^\ddagger_{m,q}, K^\ddagger) = \mathrm{GL}(m,q) \tag{1.4}$$

and in the very generic case we have

$$\mathrm{Gal}(E^\sharp_{m,q}, K^\sharp) = \mathrm{GL}(m,q), \tag{1.5}$$

where for deriving these two consequences we have to take $m > 1$, but for $m = 1$ they can be easily proved directly. Note that (1.5) was originally proved by Moore in his 1896 paper [Mor]; moreover, a very simple proof of Moore's theorem (1.5) can be found in [AbS]. From (1.5) we deduce that in the binomial case we have

$$\mathrm{Gal}(E^\flat_{m,q^n}, K^\flat) = \mathrm{GL}(1, q^{mn}) \text{ provided } \mathrm{GF}(q^{mn}) \subset k_q \tag{1.6}$$

which is also obvious. Likewise, from (1.3) we deduce that in the generic case of type J^* we have

$$J^\dagger \subset J^* \Rightarrow \mathrm{Gal}(E^*_{m,q^n}, K^*) = \mathrm{GL}(m,q^n) \text{ provided } \mathrm{GF}(q^n) \subset k_q. \tag{1.7}$$

Now for the generic case of type J^* we get

$$J^\ddagger \subset J^* \Rightarrow \mathrm{Gal}(E^*_{m,q^n}, K^*) = \mathrm{GL}(m,q^n) \text{ provided } \mathrm{GF}(q^n) \subset k_q, \tag{1.8}$$

which we deduce from (1.6) or (1.7) depending on whether $m = 1$ or $m > 1$.

Without the provisos in (1.6) to (1.8), the Galois groups get bloated from GL to ΓL. More precisely, for any divisor δ of u let $\Gamma\mathrm{L}_\delta(m,q)$ be the unique group with $\mathrm{GL}(m,q) \lhd \Gamma\mathrm{L}_\delta(m,q) \lhd \Gamma\mathrm{L}(m,q)$, where $\lhd$ denotes normal subgroup, such that $\Gamma\mathrm{L}_\delta(m,q)/\mathrm{GL}(m,q) = Z_\delta$ = the cyclic group of order δ, and for any positive integer v let $\delta(k_q, v)$ be the unique divisor of v such that $\mathrm{Gal}(Y^{q^v} - Y, k_q) = Z_{\delta(k_q,v)}$. Then by the results of [Ab6] we see that for the generic polynomial of type J^* we have

$$\mathrm{Gal}(E^\flat_{m,q^n}, K^\flat) = \Gamma\mathrm{L}_{\delta(k_q,mn)}(1, q^{mn}) \tag{1.9}$$

and

$$J^{\dagger} \subset J^{*} \Rightarrow \mathrm{Gal}(E^{*}_{m,q^n}, K^{*}) = \Gamma \mathrm{L}_{\delta(k_q,n)}(m, q^n) \tag{1.10}$$

and

$$J^{\ddagger} \subset J^{*} \Rightarrow \mathrm{Gal}(E^{*}_{m,q^n}, K^{*}) = \Gamma \mathrm{L}_{\delta(k_q,n)}(m, q^n). \tag{1.11}$$

To mitigate this bloating we take recourse to generalized iteration as defined below.

DEFINITION (1.12). For every nonnegative integer j we inductively define the jth *iterate* $E^{[j]}$ of E by putting $E^{[0]} = E^{[0]}(Y) = Y, E^{[1]} = E^{[1]}(Y) = E(Y)$, and $E^{[j]} = E^{[j]}(Y) = E(E^{[j-1]}(Y))$ for all $j > 1$. Next we define the *generalized* rth *iterate* $E^{[r]}$ of E for any $r = r(T) = \sum_i r_i T^i \in \Omega[T]$ with $r_i \in \Omega$ (and $r_i = 0$ for all except a finite number of i), where T is an indeterminate, by putting $E^{[r]} = E^{[r]}(Y) = \sum_i E^{[i]}(Y)$. Note that, for the Y-derivative $E^{[r]}_Y(Y)$ of $E^{[r]}(Y)$ we clearly have

$$E^{[r]}_Y(Y) = E^{[r]}_Y(0) = r(X_m), \tag{1.13}$$

and hence if $r(X_m) \neq 0$ then $E^{[r]}$ is a separable vectorial q-polynomial over Ω whose q-degree in Y equals m times the T-degree of r. Also note that the definition of $E^{[r]}$ remains valid for any vectorial E without assuming it to be monic or separable. Moreover, in such a general set-up, this makes the additive group of all vectorial q-polynomials $E = E(Y)$ in Y over Ω into a $\Omega[T]$-*premodule* having all the properties of a module except the left distributive law and the associativity of the multiplication; i.e., for all $r, r' \in \Omega[T]$ we have $E^{[r+r']} = E^{[r]} + E^{[r']}$, but and for all E, E' over Ω we need not have $(E + E')^{[r]} = E^{[r]} + E'^{[r]}$, and in general $E^{[rr']}$ need not be equal to $(E^{[r]})^{[r']}$. At any rate, $E^{[j]}$ of the previous notation corresponds to $E^{[T^j]}$ in the present notation. Reverting to the fixed monic separable vectorial E exhibited in (1.1), the said premodule structure makes Ω into a $\mathrm{GF}(q)[T]$-module when for every $r \in \mathrm{GF}(q)[T]$ and $z \in \Omega$ we define the "product" of r and z to be $E^{[r]}(z)$; we denote this $\mathrm{GF}(q)[T]$-module by Ω_E. Now let us fix

$$s = s(T) \in R = \mathrm{GF}(q)[T] \text{ of } T\text{-degree } n \text{ with } s(X_m) \neq 0 \tag{1.14}$$

and note that then $E^{[s]}$ is a separable vectorial q-polynomial of q-degree mn in Y over K, and the coefficient of its highest degree term equals the coefficient of the highest degree of $s(T)$. Let $V^{[s]}$ be the set of all roots of $E^{[s]}$ in Ω, and note that then $V^{[s]}$ is an (mn)-dimensional $\mathrm{GF}(q)$-vector-subspace of Ω. Let $\mathrm{GF}(q, s) = R/sR$, where sR is the ideal generated by s in $R = \mathrm{GF}(q)[T]$, and let $\omega : R \to \mathrm{GF}(q, s)$ be the canonical epimorphism. Now $V^{[s]}$ is a submodule

of Ω_E and as such it is annihilated by sR and hence we may regard it as a $\mathrm{GF}(q,s)$-module; note that then, for every $r \in R$ and $z \in \Omega$, the "product" of $\omega(r)$ and z is given by $\omega(r)z = E^{[r]}(z) = \sum_i E^{[i]}(z)$, and for every $g \in \mathrm{Gal}(K(V^{[s]}), K)$ we have $g(\omega(r)z) = \sum g(r_i) E^{[i]}(g(z)) = (\omega(r))g(z)$. It follows that, in a natural manner,

$$\mathrm{Gal}(E^{[s]}, K) < \mathrm{GL}(V^{[s]}), \tag{1.15}$$

where $\mathrm{GL}(V^{[s]})$ is the *group of all* $\mathrm{GF}(q,s)$*-linear automorphisms* of $V^{[s]}$, by which we mean all additive isomorphisms $\sigma: V^{[s]} \to V^{[s]}$ such that for all $\eta \in \mathrm{GF}(q,s)$ and $z \in V^{[s]}$ we have $\sigma(\eta z) = \eta\sigma(z)$. Note that

$$s \text{ irreducible in } R \Rightarrow \mathrm{GL}(V^{[s]}) \approx \mathrm{GL}(m, q^n) \tag{1.16}$$

where $\approx$ denotes isomorphism.

Note (1.17). The above definition is a paraphrase of Remark (3.30) of [Ab7]. As corrections to that remark: (1) The last half of the fourth sentence of that Remark, which reads "except additivity … need not have $E^{[r+r']} = E^{[r]} + E^{[r']}$" should be changed to the last half of the fifth sentence of the above definition, which reads "except the left … need not be equal to $(E^{[r]})^{[r']}$." (2) The last half of the sixth sentence of that remark, which now reads "makes Ω into a $\Omega[T]$-module … we denote it by Ω_E" should be changed to the last half of the seventh sentence of the above definition (1.12), which reads "makes Ω into a $\mathrm{GF}(q)[T]$-module … we denote this $\mathrm{GF}(q)[T]$-module by Ω_E." (3) In the 11th sentence of the same remark, the phrase "$r \in \Omega[T]$" should be replaced by the phrase "$r \in \mathrm{GF}(q)[T]$."

In Note (3.37) of [Ab7] we asked if in the very generic case we have $\mathrm{Gal}(E^{\sharp[s]}, K^{\sharp}) = \mathrm{GL}(V^{[s]})$. In [AbS] we proved this when $s = T^n$ and in Theorem (3.25) of [Ab7] we semilinearized that result. The aim of the present paper is to prove it when s is irreducible under the assumptions that m is a square-free (i.e., not divisible by the square of any prime) integer with $\mathrm{GCD}(m,n) = 1$, and $\mathrm{GCD}(mnu, 2p) = 1$. Actually, as a slightly more general result, we have the following:

MAIN THEOREM (1.18). *Assume that s is irreducible in R, and $J^{\ddagger} \subset J^*$. Also assume that m is a square-free integer with $GCD(m,n) = 1$, and $GCD(mnu, 2p) = 1$. Then in the generic case of type J^** we have $Gal(E^{*[s]}_{m,q}, K^*) = GL(V^{[s]}) \approx GL(m, q^n)$.

Now CPT (= the Classification of Projectively Transitive Permutation Groups, i.e., subgroups of GL acting transitively on nonzero vectors) is a remarkable consequence of CT (= the Classification Theorem of Finite Simple Groups). The implication CT $\Rightarrow$ CPT was mostly proved by Hering

[He1, He2]; it is also discussed by Cameron [Cam], Kantor [Kan], and Liebeck [Lie]. The proof of the Main Theorem, to be given in Section 2, will make essential use of the following weaker version of CPT.

Weak CPT (1.19). Let d be an odd positive integer, and let $G < GL(d, p)$ be transitive on the nonzero vectors $GF(p)^d \backslash \{0\}$. Then there exist positive integers b, c with $bc = d$ and a group G_0 with $SL(b, p^c) < G_0 < \Gamma L(b, p^c)$ such that $G \approx G_0$.

The $m = 1$ case of (1.18), without the hypothesis $\mathrm{GCD}(mnu, 2p) = 1$ but assuming $k_q = \mathrm{GF}(q)$, was proved by Carlitz [Car] in connection with his explicit class field theory. In our proof of the Main Theorem we need the following variation (1.20) of Carlitz's result and, for the sake of completeness, in Section 3 we shall give a self-contained proof of it including Hayes' [Hay] version of Carlitz's result. Specifically, the $m = 1$ case of (1.20) will be proved in (3.1), and the $\mathrm{GCD}(m, n) = 1$ case of (1.20) will be proved in (3.2). Recall that a univariate polynomial $\tilde{F}(Y) = \sum_{i=0}^{N} \tilde{F}_i Y^i$ of positive degree N in Y is said to be *Eisenstein* relative to $(\tilde{R}, \tilde{M})$, where $\tilde{M}$ is a prime ideal in a ring $\tilde{R}$, if $\tilde{F}_N \in \tilde{R} \backslash \tilde{M}$, $\tilde{F}_i \in \tilde{M}$ for $1 \leq i \leq N - 1$, and $\tilde{F}_0 \in \tilde{M} \backslash \tilde{M}^2$.

Carlitz Irreducibility Lemma (1.20). *Assume that s is irreducible in R, and $J^p \subset J^*$. Let $s^*(T)$ be a nonconstant irreducible factor of $s(T)$ in $k_q[T]$, and let M^* be the ideal in $R^* = k_q[X_{J_1^*}, \ldots, X_{J_{W^*}^*}]$ generated by $X_{J_1^*}, \ldots, X_{J_{W^*-1}^*}, s^*(X_m)$. Then, for $m = 1$, in the generic case of type J^* we have that $M^* = s^*(X_m)R^*$ is a maximal ideal in $R^* = k_q[X_m]$, $Y^{-1}E_{1,q}^{*[s]}(Y)$ is Eisenstein relative to (R^*, M^*), $Y^{-1}E_{1,q}^{*[s]}(Y)$ is irreducible in $K^*[Y]$, and $Gal(E_{1,q}^{*[s]}, K^*) = GL(V^{[s]}) \approx GL(1, q^n)$. Moreover, without assuming $m = 1$, but assuming $GCD(m, n) = 1$, in the generic case of type J^* we have that M^* is a maximal ideal in R^*, $Y^{-1}E_{m,q}^{*[s]}(Y)$ is Eisenstein relative to (R^*, M^*), $Y^{-1}E_{m,q}^{*[s]}(Y)$ is irreducible in $K^*[Y]$, and $Gal(E_{m,q}^{*[s]}, K^*)$ has an element of order $q^{mn} - 1$.*

The following number-theoretic result was originally proved in Zsigmondy's 1892 paper [Zsi]; proofs can also be found in Birkhoff and Vandiver's 1904 paper [BVa], Dickson's 1905 paper [Dic], Artin's 1955 paper [Art], and Feit's 1988 paper [Fei].

Zsigmondy's Theorem (1.21). *Let $M > 1$ and $N > 1$ be any integers. Assume that $(M, N) \neq (2, 6)$. Also assume that $(M, N) \neq (2^i - 1, 2)$ for any integer $i > 1$, (recall that a prime number of the form $2^i - 1$ is called a Mersenne prime, and in that case i is automatically a prime). Then $M^N - 1$ has a prime divisor which does not divide $M^{N'} - 1$ for any positive integer $N' < N$.*

In Section 4 we shall prove the following consequence of (1.21), needed in the proof of the Main Theorem.

Order Divisibility Lemma (1.22). *Let b, c, l be any positive integers with b odd such that $bc = ml$ and $|SL(b, p^c)|$ divides $|GL(m, p^l)|$. Then b divides m.*

In proving our Main Theorem, in addition to items (1.19), (1.20), and (1.22), we shall also use the first part of the following well-known versatile lemma initiated by Singer [Sin] (cf. pp. 34 and 105 of Dembowski [Dem], pp. 84–91 of Lidl and Niederreiter [LNi], and pp. 187–189 of Huppert [Hup]):

SINGER CYCLE LEMMA. (1.23) *Let* $A \in GL(m,q)$ *have order* $e = q^m - 1$. *Then* $\det(A)$ *has order* $\varepsilon = q - 1$, *and* A *acts transitively on the nonzero vectors* $GF(q)^m \setminus \{0\}$; *i.e., it is an e-cycle in the symmetric group* S_e (*and as such it is called a Singer Cycle*). *Moreover, in* $GL(m,q)$ *all subgroups generated by such elements, i.e., all cyclic subgroups of order e, form a nonempty complete set of conjugates.*

For the convenience of the reader (which really means for our own benefit), in (5.12) of Section 5 we shall give a self-contained proof of the implication $\text{ord}(A) = e \Rightarrow \text{ord}(\det(A)) = \varepsilon$, which is the only part of (1.23) we shall use in the proof of our Main Theorem.

Some parts of the proof of our Main Theorem could have been deduced from Drinfeld module theory (cf. the recent book [Gos] of David Goss). However, to make our paper self-contained, we do not fall back on that theory.

It is a great pleasure to thank Paul Eakin, Walter Feit, Pradipkumar Keskar, Paul Loomis, Shashikant Mulay, Michael O'Nan, Avinash Sathaye, and John Thompson for many stimulating conversations regarding the material of this paper.

2. MAIN THEOREM

Assume that we are in the generic case of type J^*.
For a moment

$$\begin{gathered}\text{suppose that } W^* > 0 \text{ and } j < m \text{ where } j = J_1^*; \\ \text{also let } X = X_j \text{ and } R^* = k_q[X_{J_1^*}, \dots, X_{J_{W^*}^*}].\end{gathered} \tag{$*$}$$

By induction we shall show that for any positive integer v we have

$$(*) \Rightarrow \begin{cases} E_{m,q}^{*[T^v]}(Y) = Y^{q^{mv}} + \sum_{i=j}^{mv} D_{v,i} Y^{q^{mv-i}} \\ \text{where } D_{v,i} \in R^* \text{ with } D_{v,j} = \sum_{\lambda=0}^{v-1} X^{q^{m\lambda}}. \end{cases} \tag{2.1*}$$

This is obvious for $v = 1$ because $E_{m,q}^{*[T]}(Y) = E_{m,q}^*(Y)$ and hence

$$E_{m,q}^{*[T]}(Y) = Y^{q^m} + \sum_{w=j}^{m} X_w Y^{q^{m-w}}$$

where we recall that $X_w = 0$ for $w \notin J^*$, and $X_j = X$. So let $v > 1$ and assume true for $v - 1$. Then

$$E_{m,q}^{*[T^{v-1}]}(Y) = Y^{q^{mv-m}} + \sum_{i=j}^{mv-m} D_{v-1,i} Y^{q^{mv-m-i}}$$

$$\text{where } D_{v-1,i} \in R^* \text{ with } D_{v-1,j} = \sum_{\lambda=0}^{v-2} X^{q^{m\lambda}}$$

and substituting this for Y in the above expression for $E_{m,q}^{*[T]}(Y)$ we get

$$\begin{aligned} E_{m,q}^{*[T^v]}(Y) &= \left(Y^{q^{mv-m}} + \sum_{i=j}^{mv-m} D_{v-1,i} Y^{q^{mv-m-i}} \right)^{q^m} \\ &\quad + \sum_{w=j}^{m} X_w \left(Y^{q^{mv-m}} + \sum_{i=j}^{mv-m} D_{v-1,i} Y^{q^{mv-m-i}} \right)^{q^{m-w}} \\ &= \left(Y^{q^{mv}} + \sum_{i=j}^{mv-m} D_{v-1,i}^{q^m} Y^{q^{mv-i}} \right) \\ &\quad + \left(\sum_{w=j}^{m} X_w Y^{q^{mv-w}} + \sum_{w=j}^{m} \sum_{i=j}^{mv-m} X_w D_{v-1,i}^{q^{m-w}} Y^{q^{mv-i-w}} \right) \\ &= Y^{q^{mv}} + \sum_{w=j}^{mv} D_{v,i} Y^{q^{mv-i}} \end{aligned}$$

where

$$D_{v,i} \in R^* \text{ with } D_{v,j} = X + D_{v-1,j}^{q^m} = \sum_{\lambda=0}^{v-1} X^{q^{m\lambda}}$$

which completes the induction. Now

$$s(T) = \sum_{v=0}^{n} s_v T^v \text{ where } s_v \in \mathrm{GF}(q) \text{ with } s_n \neq 0$$

and

$$E_{m,q}^{*[s]}(Y) = \sum_{v=0}^{n} s_v E_{m,q}^{*[T^v]}$$

and hence by (2.1*) we see that

$$(*) \Rightarrow \begin{cases} E^{*[s]}_{m,q}(Y) = Y^{q^{mn}} + \sum_{i=j}^{mn} D_{s,i} Y^{q^{mn-i}} & \text{where} \\ D_{s,i} \in R^* \text{ with } D_{s,j} = s_n \sum_{\lambda=0}^{n-1} X^{q^{m\lambda}} \text{ and } 0 \neq s_n \in \mathrm{GF}(q). \end{cases} \tag{2.2*}$$

Continuing with supposition (*), also suppose that $J^\flat \subset J^*$. Let k be the algebraic closure of k_q in Ω. Then we can take elements $(x_{J^*_2}, \ldots, x_{J^*_{W^*}})$ in k such that $s(x_{J^*_{W^*}}) \neq 0$. Let $F^{[s]} = F^{[s]}(Y)$ be the polynomial obtained by specializing $E^{*[s]}_{m,q}(Y)$ at $(X_{J^*_2}, \ldots, X_{J^*_{W^*}}) = (x_{J^*_2}, \ldots, x_{J^*_{W^*}})$. Clearly $F^{[s]}(Y)$ is a vectorial q-polynomial of q-degree mn in Y over $k(X)$. Now the Y-derivative of $F^{[s]}(Y)$ is $s(x_{J^*_{W^*}})$ which is a nonzero element of k, and hence $F^{[s]}(Y)$ is separable and its Y-discriminant is a nonzero element of k. Therefore $F^{[s]}(Y)$ gives an unramified covering of the affine line over k, and hence $\mathrm{Gal}(F^{[s]}, k(X))$ is a quasi-p group, i.e., it is generated by all of its p-Sylow subgroups (cf. [Ab1]). By (2.2*) we see that the coefficient of $Y^{q^{mn-j}}$ in $F^{[s]}(Y)$ is a nonconstant polynomial in X. Therefore the polynomial $Y^{-1}F^{[s]}(Y)$ belongs to $k[X][Y] \backslash k[Y]$. Moreover, as a polynomial in Y, its highest degree coefficient and its constant term are both nonzero elements of k. Hence, by the Gauss lemma, the said polynomial cannot factor completely into linear factors in Y over $k(X)$. Consequently $\mathrm{Gal}(F^{[s]}, k(X)) \neq \{1\}$, and hence p divides $|\mathrm{Gal}(F^{[s]}, k(X))|$. Since by the above specialization, $\mathrm{Gal}(F^{[s]}, k(X))$ can be identified with a subgroup of $\mathrm{Gal}(E^{*[s]}_{m,q}, K^*)$, we see that p divides $|\mathrm{Gal}(E^{*[s]}_{m,q}, K^*)|$. Thus

$$J^\flat \subset J^* \text{ with } J^\flat \neq J^* \Rightarrow p \text{ divides } |\mathrm{Gal}(E^{*[s]}_{m,q}, K^*)|. \tag{2.3}$$

To prove the Main Theorem, henceforth in this section assume that s is irreducible in R and $J^\ddagger \subset J^*$; also assume that m is a square-free integer with $\mathrm{GCD}(m, n) = 1$ and $\mathrm{GCD}(mnu, 2p) = 1$. By (1.15), (1.16) and (1.20) we see that

$$m = 1 \Rightarrow \mathrm{Gal}(E^{*[s]}_{m,q}, K^*) = \mathrm{GL}(V^{[s]}) \approx \mathrm{GL}(m, q^n) \tag{2.4}$$

and therefore henceforth we may also assume that $m > 1$. Again by (1.15) and (1.16) we see that

$$\mathrm{Gal}(E^{*[s]}_{m,q}, K^*) < \mathrm{GL}(V^{[s]}) \approx \mathrm{GL}(m, q^n). \tag{2.5}$$

By (1.20) we know that $Y^{-1}E^{*[s]}_{m,q}(Y)$ is irreducible in $K^*[Y]$ and clearly $E^{*[s]}_{m,q}(Y)$ is a separable vectorial p-polynomial of p-degree mnu in Y over K^*, and hence $\mathrm{Gal}(E^{*[s]}_{m,q}, K^*)$ is isomorphic to a subgroup of $\mathrm{GL}(d, p)$, where $d = mnu$, acting transitively on the nonzero vectors $\mathrm{GF}(p)^d \backslash \{0\}$; therefore by

(1.19) we see that there exist positive integers b, c with $bc = d$ and a group G_0 such that

$$\mathrm{SL}(b, p^c) < G_0 < \Gamma\mathrm{L}(b, p^c) \text{ and } \mathrm{Gal}(E^{*[s]}_{m,q}, K^*) \approx G_0. \tag{2.6}$$

By assumption $\mathrm{GCD}(d, 2p) = 1$, and hence p does not divide $|\Gamma\mathrm{L}(1, p^d)|$; but by (2.3) we know that p divides $|\mathrm{Gal}(E^{*[s]}_{m,q}, K^*)|$; therefore by (2.6) we must have $b \neq 1$. By (2.5) and (2.6) we know that $|SL(b, p^c)|$ divides $|\mathrm{GL}(m, p^{nu})|$, and hence in view of (1.22), by choosing $l = nu$, we see that b divides m. Thus

$$b > 1 \text{ and } b \text{ divides } m. \tag{2.7}$$

For a moment suppose that m is prime. Then by (2.7) we must have $b = m$. Therefore by (2.5) and (2.6) we see that for some group $G_0^\sharp$ we have

$$\mathrm{SL}(m, p^{nu}) < G_0 \approx G_0^\sharp < \mathrm{GL}(m, p^{nu}). \tag{2.8′}$$

Using the fact that SL is generated by elements of p-power order of GL, from (2.8′) we deduce (see footnotes 13 and 14 of [Ab6] characterizing the quasi-p part of a finite group) the fact that for the group $G_0^\sharp$ mentioned in (2.8′) we have

$$\mathrm{SL}(m, p^{nu}) < G_0^\sharp < \mathrm{GL}(m, p^{nu}) \quad \text{and} \quad G_0 \approx G_0^\sharp. \tag{2.9′}$$

In view of (2.6) and (2.8′), by (1.20) we can find $A_0 \in G_0^\sharp$ with $\mathrm{ord}(A_0) = e = p^{mnu} - 1$. Now by (1.23) we get $\mathrm{ord}(\det(A_0)) = \varepsilon = p^{nu} - 1$ and hence the image of A_0 under the determinant map of $\mathrm{GL}(m, p^{nu})$ onto $\mathrm{GF}(p^{nu})^*$ generates all of $\mathrm{GF}(p^{nu})^*$, and hence the image of $G_0^\sharp$ under the said map coincides with $\mathrm{GF}(p^{nu})^*$. Since the kernel of the said map is $\mathrm{SL}(m, p^{nu})$, which by (2.9′) is contained in $G_0^\sharp$, we must have $G_0^\sharp = \mathrm{GL}(m, p^{nu})$, and hence by (2.5), (2.6), and (2.8′) we get $\mathrm{Gal}(E^{*[s]}_{m,q}, K^*) = \mathrm{GL}(V^{[s]}) \approx \mathrm{GL}(m, q^n)$. Thus

$$m \text{ is prime} \Rightarrow \mathrm{Gal}(E^{*[s]}_{m,q}, K^*) = \mathrm{GL}(V^{[s]}) \approx \mathrm{GL}(m, q^n). \tag{2.10}$$

Finally, suppose that $m > 1$ and m is not prime. Let $m = m_1 m_2 \dots m_\tau$, where $m_1, m_2, \dots, m_\tau$ are distinct primes. For $1 \leq t \leq \tau$, let $q_t = q^{m/m_t}$, let J_t be the subsequence of J^* consisting of those $i \in J^*$ which are divisible by m/m_t, and let $F_t = F_t(Y)$ and K_t be the respective specializations of $E^*_{m,q}(Y)$ and K^* obtained by putting $X_i = 0$ for i not divisible by m/m_t and relabelling the remaining X_i as $X_{im_t/m}$. Then $F_t^{[s]}(Y)$ is the corresponding specialization of $E^{*[s]}_{m,q}(Y)$ and, since $\mathrm{GCD}(m, n) = 1$, the polynomial s remains irreducible in

$\mathrm{GF}(q_t)[T]$. Clearly $F_t(Y) = E^{**}_{m_t,q_t}(Y)$ where the sequence J^{**} is obtained from the sequence J_t by changing every i in J_t to im_t/m, and hence by (2.10) we see that if $\mathrm{GF}(q_t) \subset k_q$ then $\mathrm{Gal}(F_t^{[s]}, K_t) \approx \mathrm{GL}(m_t, q_t^n)$. Therefore, by the basic extension principle on page 93 of [Ab2] (which says that extending the ground field reduces the Galois group to a subgroup), and by the fact that $F_t^{[s]}(Y)$ is a specialization of $E^{*[s]}_{m,q}(Y)$, we conclude that $\mathrm{GL}(m_t, q_t^n)$ is isomorphic to a subgroup of $\mathrm{Gal}(E^{*[s]}_{m,q}, K^*)$. Consequently, since $\mathrm{SL}(m_t, q_t^n) < \mathrm{GL}(m_t, q_t^n)$, we see that $\mathrm{SL}(m_t, q_t^n)$ is isomorphic to a subgroup of $\mathrm{Gal}(E^{*[s]}_{m,q}, K^*)$. Therefore we can find a group H_t such that

$$H_t < \mathrm{Gal}(E^{*[s]}_{m,q}, K^*) \quad \text{and} \quad H_t \approx \mathrm{SL}(m_t, q_t^n). \tag{2.11''}$$

(Note that we numbered the above items as (2.8′) and (2.9′) to indicate that there we are in the case of prime m; likewise we are numbering the current items as (2.11″) to (2.17″) to indicate that they pertain to the case of nonprime $m > 1$.) By (2.6) and (2.11″) we see that for some group H_0 we have

$$H_t \approx \mathrm{SL}(m_t, q_t^n) \approx H_0 < G_0 < \Gamma L(b, p^c). \tag{2.12''}$$

In particular, by (2.12″) we get

$$\mathrm{SL}(m_t, q_t^n) \approx H_0 < \Gamma\mathrm{L}(b, p^c). \tag{2.13''}$$

Since $m_t > 2$, upon letting prime denote commutator subgroup, we get $\mathrm{SL}(m_t, q_t^n)' = \mathrm{SL}(m_t, q_t^n)$ (cf. Satz 6.10 on p. 181 of [Hup]). Next, since the quotient of $\Gamma\mathrm{L}(b, p^c)$ by $\mathrm{GL}(b, p^c)$ is abelian, we have $\Gamma\mathrm{L}(b, p^c)' < \mathrm{GL}(b, p^c)$. Therefore by applying primes to (2.13″) we get

$$\mathrm{SL}(m_t, q_t^n) \approx H_0 < \Gamma\mathrm{L}(b, p^c)' < \mathrm{GL}(b, p^c). \tag{2.14''}$$

Consequently, $|\mathrm{SL}(m_t, q_t^n)|$ divides $|\mathrm{GL}(b, p^c)|$, and hence by (1.22) we see that

$$m_t \text{ divides } b. \tag{2.15''}$$

Since (2.15″) is true for $1 \leq t \leq \tau$, and m is equal to the product of the distinct primes m_t, it follows that m divides b. Hence by (2.7) we must have $b = m$. Therefore by (2.5) and (2.6) we see that for some group $G^\sharp$ we have

$$\mathrm{SL}(m, p^{nu}) < G_0 \approx G^\sharp < \mathrm{GL}(m, p^{nu}). \tag{2.16''}$$

Using the fact that SL is generated by elements of p-power order of GL, from (2.16″) we deduce (see footnotes 13 and 14 of [Ab6]) the fact that for the

group $G^\sharp$ mentioned in (2.16″) we have

$$\mathrm{SL}(m, p^{nu}) < G^\sharp < \mathrm{GL}(m, p^{nu}) \text{and} G_0 \approx G^\sharp. \tag{2.17″}$$

In view of (2.6) and (2.16″), by (1.20) we can find $A \in G^\sharp$ with $\mathrm{ord}(A) = e = p^{mnu} - 1$. Now by (1.23) we get $\mathrm{ord}(\det(A)) = \varepsilon = p^{nu} - 1$ and hence the image of A under the determinant map of $\mathrm{GL}(m, p^{nu})$ onto $\mathrm{GF}(p^{nu})^*$ generates all of $\mathrm{GF}(p^{nu})^*$, and hence the image of $G^\sharp$ under the said map coincides with $\mathrm{GF}(p^{nu})^*$. Since the kernel of the said map is $\mathrm{SL}(m, p^{nu})$ which by (2.17″) is contained in $G^\sharp$, we must have $G^\sharp = \mathrm{GL}(m, p^{nu})$, and hence by (2.5), (2.6), and (2.16″) we get $\mathrm{Gal}(E^{*[s]}_{m,q}, K^*) = \mathrm{GL}(V^{[s]}) \approx \mathrm{GL}(m, q^n)$. Thus

$$m > 1 \text{ and } m \text{ is not prime} \Rightarrow \mathrm{Gal}(E^{*[s]}_{m,q}, K^*) = \mathrm{GL}(V^{[s]}) \approx \mathrm{GL}(m, q^n). \tag{2.18}$$

3. CARLITZ IRREDUCIBILITY

Assume that s is irreducible in $R = \mathrm{GF}(q)[T]$, and we are in the generic case of type J^* with $J^\flat \subset J^*$. Let $s^*(T)$ be a nonconstant irreducible factor of $s(T)$ in $k_q[T]$, and let M^* be the ideal in $R^* = k_q[X_{J_1^*}, \ldots, X_{J_{W^*}^*}]$ generated by $X_{J_1^*}, \ldots, X_{J_{W^*-1}^*}, s^*(X_m)$. Then obviously M^* is a maximal ideal in R^*.

For a moment suppose that $m = 1$ and, in addition to using the notation of Definition (1.12) with $E(Y) = E^*_{1,q}(Y)$ and $K = K^*$, let us put $\hat{R} = \mathrm{GF}(q)[X_m]$ and $\hat{K} = \mathrm{GF}(q)(X_m)$, and note that then $\mathrm{GF}(q) \subset k_q$ and hence $\hat{R} \subset R^* = k_q[X_m]$ and $\hat{K} \subset K^* = k_q(X_m)$. Recall that $V^{[s]}$ is the set of all roots of $E^{[s]}(Y)$ in Ω. Now $V^{[s]}$ is an n-dimensional vector space over $\mathrm{GF}(q)$ and hence it is a 1-dimensional vector space over $\mathrm{GF}(q^n) \approx R/sR$; therefore $V^{[s]}$ is a cyclic R-module and any nonzero element of it is a generator. Let ζ be a generator of the cyclic R-module $V^{[s]}$, let R_0 be the set of all nonzero members of R of T-degree less than n, let $\hat{R}^{[s]}$ be the integral closure of $\hat{R}$ in $\hat{K}(V^{[s]})$, and let s_n be the coefficient of T^n in $s(T)$. Then $0 \neq s_n \in \mathrm{GF}(q)$ and

$$s_n^{-1} s(X_m) = s_n^{-1} E_Y^{[s]}(0) = (-1)^{q^n - 1} \prod_{r \in R_0} \omega(r)\zeta = (-1)^{q^n - 1} \prod_{r \in R_0} E^{[r]}(\zeta).$$

Now ζ divides $\omega(r)\zeta$ in $\hat{R}^{[s]}$ because Y divides $E^{[r]}(Y)$ in $\hat{R}[Y]$. Since any nonzero element of $V^{[s]}$ can be chosen as an R-module generator, by symmetry $\omega(r)\zeta$ divides ζ in $\hat{R}^{[s]}$. Consequently $\omega(r)\zeta = \alpha_r \zeta$ for some unit α_r in

$\hat{R}^{[s]}$. This being so for every r in R_0, by the above product decomposition we get

$$s(X_m) = \alpha\zeta^{q^n-1}$$

where α is also a unit in $\hat{R}^{[s]}$. Let $R^{*[s]}$ be the integral closure of R^* in K^*. Then $\hat{R}^{[s]} \subset R^{*[s]}$, and hence $\zeta \in R^{*[s]}$ and α is a unit in $R^{*[s]}$. Since GF(q) is perfect and $s(X_m)$ is irreducible in $\hat{R}$, it follows that

$$s(X_m) = s^*(X_m)s^{**}(X_m)$$

where $s^{**}(X_m) \in R^*$ is coprime with $s^*(X_m)$, and hence upon letting v to be the real discrete valuation of K^* corresponding to the maximal ideal (now a principal prime ideal) $M^* = s^*(X_m)R^*$ in R^* we have $v(\hat{s}(X_m)) = 0$. Since $\mathrm{Gal}(E^{[s]}, K^*) < \mathrm{GL}(V^{[s]}) \approx \mathrm{GL}(1, q^n) =$ a cyclic group of order $q^n - 1$, by looking at extensions of v to $K^*(V^{[s]})$, in view of the above two equations for $s(X_m)$ we see that $\mathrm{Gal}(E^{[s]}, K^*) = \mathrm{GL}(V^{[s]}) \approx \mathrm{GL}(1, q^n)$, and v is totally ramified in $K^*(V^{[s]})$, i.e., it has only one extension w to $K^*(V^{[s]})$ and the ramification exponent of w over v is $[K^*(V^{[s]}) : K^*] = q^n - 1$. Since the Y-degree of $Y^{-1}E^{[s]}(Y)$ is $q^n - 1$, and since every subgroup of a finite cyclic group is cyclic with order dividing the order of the group, it follows that $Y^{-1}E^{[s]}(Y)$ is irreducible in $K^*[Y]$ and $K^*(V^{[s]}) = K^*(\zeta)$ with $E^{[s]}(\zeta) = 0$. Since the constant term of $Y^{-1}E^{[s]}(Y)$ has v-value 1 and the leading coefficient of $Y^{-1}E^{[s]}(Y)$ has v-value 0, and since v has a unique extension to $K^*(\zeta)$, it follows that all the other coefficients of $Y^{-1}E^{[s]}(Y)$ have positive v-values, and $Y^{-1}E^{[s]}(Y)$ is Eisenstein relative to (R^*, M^*). Thus

$$m = 1 \Rightarrow \begin{cases} Y^{-1}E^{*[s]}_{1,q}(Y) \text{ is Eisenstein relative to } (R^*, M^*), \\ Y^{-1}E^{*[s]}_{1,q}(Y) \text{ is irreducible in } K^*[Y], \text{ and} \\ \mathrm{Gal}(E^{*[s]}_{1,q}, K^*) = \mathrm{GL}(V^{[s]}) \approx \mathrm{GL}(1, q^n) \\ \text{which is acyclic group of order } q^n - 1. \end{cases} \tag{3.1}$$

Next, instead of letting $m = 1$, assume that $\mathrm{GCD}(m, n) = 1$. Then the polynomial s remains irreducible in the ring $\mathrm{GF}(q^m)[T]$. Also clearly $E^{\flat}_{m,q}(Y)$ is the specialization of $E^*_{m,q}(Y)$ obtained by putting $X_i = 0$ for $i < m$, and $E^{\flat[s]}_{m,q}(Y)$ is the corresponding specialization of $E^{*[s]}_{m,q}(Y)$, and hence $\mathrm{Gal}(E^{\flat[s]}_{m,q}, K^*)$ is isomorphic to a subgroup of $\mathrm{Gal}(E^{*[s]}_{m,q}, K^*)$. By putting $X_m = X_1$ we get $E^{\flat}_{m,q}(Y) = E^{\flat}_{1,q^m}(Y)$ and hence we get $E^{\flat[s]}_{m,q}(Y) = E^{\flat[s]}_{1,q^m}(Y)$, and therefore by (3.1) we see that $Y^{-1}E^{*[s]}_{m,q}(Y)$ is Eisenstein relative to (R^*, M^*), $Y^{-1}E^{*[s]}_{m,q}(Y)$ is irreducible in $K^*[Y]$, and $\mathrm{Gal}(E^{*[s]}_{m,q}, K^*)$ has an

element of order $q^{mn} - 1$. Thus

$$(m,n) = 1 \Rightarrow \begin{cases} Y^{-1}E^{*[s]}_{m,q}(Y) \text{ is Eisenstein relative to } (R^*, M^*), \\ Y^{-1}E^{*[s]}_{m,q}(Y) \text{ is irreducible in } K^*[Y], \text{ and} \\ \mathrm{Gal}(E^{*[s]}_{m,q}, K^*) \text{ has an element of order } q^{mn} - 1. \end{cases} \tag{3.2}$$

4. ORDER DIVISIBILITY

To deduce the Order Divisibility Lemma from Zsigmondy's Theorem, let b, c, d, l be any positive integers with b odd such that $bc = d = ml$ and $|\mathrm{SL}(b, p^c)|$ divides $|\mathrm{GL}(m, p^l)|$. We want to show that then b divides m. If $b = 1$ then we have nothing to show. So assume that $b \geq 3$. By the standard order formulas we have

$$|\mathrm{SL}(b, p^c)| = p^{d(b-1)/2} \prod_{0 \leq j \leq b-2} (p^{d-jc} - 1)$$

and

$$|\mathrm{GL}(m, p^l)| = p^{d(m-1)/2} \prod_{0 \leq j \leq m-1} (p^{d-jl} - 1).$$

Since by assumption the order of the SL divides the order of the GL, and since the zeroth terms in the two products coincide, the first term of the product in the SL formula must divide the GL formula after we delete from it the said zeroth term. In other words

$$(p^{d-c} - 1) \text{ divides } p^{d(m-1)/2} \prod_{1 \leq j \leq m-1} (p^{d-jl} - 1). \tag{4.1}$$

Again, since by assumption the order of the SL divides the order of the GL, and since the terms in the two products are all nondivisible by p, comparing the exponents of the p-parts we get $d(b-1)/2 \leq d(m-1)/2$ and hence $b \leq m$. Since $bc = d$ and b is odd with $b \geq 3$, we see that if $d - c \leq 6$ then $(b, c) = (3,1)$ or $(3,2)$ or $(3,3)$ or $(5,1)$ or $(7,1)$; since $bc = ml$ and $b \leq m$, in each of these cases b divides m. So assume that $d - c > 6$. Then by Zsigmondy's Theorem, we can find a prime π such that

$$\pi \text{ divides } p^{d-c} - 1 \tag{4.2}$$

but

$$\pi \text{ does not divide } p^i - 1 \text{ for any positive integer } i < d - c. \tag{4.3}$$

By (4.1) and (4.2), there exist an integer j such that

$$\pi \text{ divides } p^{d-jl} - 1 \text{ and } 1 \leq j \leq m-1. \tag{4.4}$$

By (4.3) and (4.4), we get $d - jl \geq d - c$ and hence

$$c \geq jl. \tag{4.5}$$

Now let us make a general observation about Zsigmondy primes, i.e., about primes π which divide $M^N - 1$ but not $M^{N'} - 1$ for any positive integer $N' < N$; here M and N are positive integers. Let N'' be any positive integer such that π divides $M^{N''} - 1$. Upon letting M denote the image of M under the residue class epimorphism of the ring of integers onto $\mathrm{GF}(\pi)$ we get $\bar{M}^N = 1 = \bar{M}^{N''}$ and hence by order consideration we must have $\bar{M}^{N^*} = 1$ where $\mathrm{GCD}(N, N'') = N^*$; therefore by the characterizing property of Zsigmondy primes we get $N^* \geq N$ and this together with the fact that $\mathrm{GCD}(N, N'') = N^*$ implies that N divides N''. Applying this to our situation, by (4.2) to (4.4) we conclude that $d - c$ divides $d - jl$; consequently, by subtracting the first number from the second, we see that $d - c$ divides $c - jl$. Remembering that $d = bc$, we conclude that $c(b-1)$ divides $c - jl$; therefore

$$c \text{ divides } c - jl. \tag{4.6}$$

By (4.5) and (4.6) we get $c = jl$, and hence l divides c; since $bc = ml$, it follows that b divides m.

5. SINGER CYCLES

Recall that the multiplicative group $\mathrm{GF}(q^m)^*$ of all nonzero elements of $\mathrm{GF}(q^m)$ is a cyclic group of order $q^m - 1$, and a generator of this group is called a primitive element of $\mathrm{GF}(q^m)$. Note that a primitive element of $\mathrm{GF}(q^m)$ is always a primitive element of $\mathrm{GF}(q^m)$ over $\mathrm{GF}(q)$ (where we regard $\mathrm{GF}(q)$ as a subfield of $\mathrm{GF}(q^m)$), but in general the converse in not true. Let $\mathrm{GF}(q^m)^{**}$ be the set of all *primitive elements* of $\mathrm{GF}(q^m)$. Let $P(m, q)$ be the set of monic polynomials

$$f = f(T) = a_0 + a_1 T + \cdots + a_{m-1} T^{m-1} + T^m$$

in an indeterminate T with coefficients $a_0, a_1, \ldots, a_{m-1}$ in $\mathrm{GF}(q)$ and $f(0) = a_0 \neq 0$. Let $P(m, q)^*$ be the set of all f in $P(m, q)$ such that f is *irreducible* in $R = GF(q)[T]$. Let $P(m, q)^{**}$ be the set of all f in $P(m, q)^*$ such that $f(U) = 0$ for some U in $\mathrm{GF}(q^m)^{**}$; members of $P(m, q)^{**}$ are called *primitive*

polynomials of degree m over $\mathrm{GF}(q)$. For any f in $P(m,q)$ and U in an overfield of $\mathrm{GF}(q)$ we have $f(U^{q^i}) = f(U)^{q^i}$ for all $i \geq 0$ and hence if $f(U) = 0$ then $f(U^{q^i}) = 0$ for all $i \geq 0$; since $U \mapsto U^{q^i}$ give distinct $\mathrm{GF}(q)$-automorphisms of $\mathrm{GF}(q^m)$ for $0 \leq i \leq m-1$, we see that

$$f \in P(m,q)^* \text{ and } f(U) = 0 \Rightarrow f(T) = \prod_{i=0}^{m-1} (T - U^{q^i}) \tag{5.1}$$

and hence $(-1)^m f(0) = U^{e/\varepsilon}$, where $e = q^m - 1$ and $\varepsilon = q - 1$, and therefore

$$f \in P(m,q)^{**} \Rightarrow (-1)^m f(0) \in \mathrm{GF}(q)^{**} \Rightarrow \mathrm{ord}((-1)^m f(0)) = \varepsilon, \tag{5.2}$$

where ord denotes order in a group. To extend the idea of order to members of $P(m,q)$, given any f in $P(m,q)$, let $\theta: R \to R/fR$ be the canonical epimorphism; then $|\theta(R)| = |\mathrm{GF}(q)^m| = q^m = e + 1$ and, because $f(0) \neq 0$, the $e+1$ elements $\theta(T^i)_{0 \leq i \leq e}$ of $\theta(R)$ are all nonzero, and hence they cannot be distinct. Consequently $\theta(T^i) = \theta(T^j)$ for some $0 \leq i < j \leq e$. Now upon letting $\rho = j - i$ we see that $1 \leq \rho \leq e$ and $T^i(T^\rho - 1)$ is divisible by $f(T)$. Again, since $f(0) \neq 0$, it follows that $T^\rho - 1$ is divisible by $f(T)$. Thus:

$$\begin{aligned} &\text{for any } f \in P(m,q) \text{ there exists a positive integer } \rho \\ &\text{such that } T^\rho - 1 \text{ is divisible by } f(T) \\ &\text{and by defining the order of } f \text{ to be the smallest such } \rho \\ &\text{and denoting it by } \mathrm{ord}(f) \text{ we have } 1 \leq \mathrm{ord}(f) \leq e. \end{aligned} \tag{5.3}$$

Since all the roots of any $f \in P(m,q)^*$ clearly have the same order and since f has no multiple roots, we see that

$$f \in P(m,q)^* \text{ and } f(U) = 0 \Rightarrow \mathrm{ord}(f) = \mathrm{ord}(U) \tag{5.4}$$

and hence in particular

$$f \in P(m,q)^{**} \Rightarrow \mathrm{ord}(f) = e. \tag{5.5}$$

Conversely, for a moment let $f \in P(m,q)$ be such that $\mathrm{ord}(f) = e$; then $f(T)$ divides $T^e - 1$ which has no multiple roots, and hence $f(T)$ has no multiple factors. Consequently, if $f \notin P(m,q)^*$ then $f(T) = g(T)h(T)$, where $g(T)$ and $h(T)$ in R are coprime monic polynomials of positive degrees b and c with $b + c = m$. Upon letting $\mathrm{ord}(g) = v$ and $\mathrm{ord}(h) = w$, by (5.3) we see that $g(T)$ divides $T^v - 1$ with $1 \leq v \leq q^b - 1$ and $h(T)$ divides $T^w - 1$ with $1 \leq w \leq q^c - 1$. Since $f(T) = g(T)h(T)$ with coprime $g(T)$ and $h(T)$, and

since $T^{vw}-1$ is divisible by T^v-1 as well as T^w-1, we see that $T^{vw}-1$ is divisible by $f(T)$; hence we get $\mathrm{ord}(f)\le vw$ which is a contradiction because $vw\le(q^b-1)(q^c-1)<(q^{b+c}-1)=e=\mathrm{ord}(f)$; therefore $f\in P(m,q)^*$ and hence by (5.4) we conclude that $f\in P(m,q)^{**}$. Thus

$$f\in P(m,q) \text{ and } \mathrm{ord}(f)=e\Rightarrow f\in P(m,q)^{**}. \tag{5.6}$$

For any $A\in \mathrm{GL}(m,q)$ let $f_A=f_A(T)$ be the *characteristic polynomial* of A; i.e.,

$$f_A(T)=\det(TI-A),$$

where I is the m by m identity matrix. Note that then $f_A(T)=\bar a_0+\bar a_1T+\cdots+\bar a_{m-1}T^{m-1}+T^m$ with $\bar a_0,\bar a_1,\ldots,\bar a_{m-1}$, in $GF(q)$, and by substituting $T=0$ in the above equation we get

$$\det(A)=(-1)^m f_A(0) \tag{5.7}$$

and hence, because $\det(A)\neq 0$, we get

$$f_A\in P(m,q). \tag{5.8}$$

By the Cayley–Hamilton Theorem we have

$$f_A(A)=0 \tag{5.9}$$

where as usual $f_A(A)=\bar a_0I+\bar a_1A+\cdots+\bar a_{m-1}A^{m-1}+A^m$. If $\mathrm{ord}(f_A)=\rho$ then $T^\rho-1=f_A(T)g(T)$ with $g(T)\in R$ and, substituting $T=A$ into this equation, by (5.9) we get $A^\rho=I$, and hence $\mathrm{ord}(A)$ divides $\mathrm{ord}(f_A)$. Therefore in view of (5.3) we see that

$$\begin{aligned}&\mathrm{ord}(f_A)/\mathrm{ord}(A)=\text{a positive integer}\\ &\text{and}\\ &1\le\mathrm{ord}(A)\le\mathrm{ord}(f_A)\le e.\end{aligned} \tag{5.10}$$

By (5.6) and (5.10) we get

$$\mathrm{ord}(A)=e\Rightarrow\mathrm{ord}(f_A)=e\Rightarrow f_A\in P(m,q) \tag{5.11}$$

and hence by (5.2) and (5.7) we see that

$$\mathrm{ord}(A)=e\Rightarrow\mathrm{ord}(\det(A))=\varepsilon. \tag{5.12}$$

REFERENCES

[Ab1] S. S. Abhyankar, Coverings of algebraic curves, *Amer. J. Math.* **79** (1957), 825–856.

[Ab2] S. S. Abhyankar, Galois theory on the line in nonzero characteristic, *Bull. Amer. Math. Soc.* **27** (1992), 68–133.

[Ab3] S. S. Abhyankar, Nice equations for nice groups, *Israel J. Math.* **88** (1994), 1–24.

[Ab4] S. S. Abhyankar, Projective polynomials, *Proc. Amer. Math. Soc.* **125** (1997), 1643–1650.

[Ab5] S. S. Abhyankar, Local fundamental groups of algebraic varieties, *Proc. Amer. Math. Soc.* **125** (1997), 1635–1641.

[Ab6] S. S. Abhyankar, Semilinear transformations, *Proc. Amer. Math. Soc.* **127** (1999), 2511–2525.

[Ab7] S. S. Abhyankar, Galois theory of semilinear transformations, *in* "Proceedings of the UF Galois Theory Week 1996" (Helmut Voelklein *et al.* (Eds.), London Mathematical Society Lecture Note Series, Vol. 256, pp. 1–37, Cambridge Univ. Press, Cambridge, UK, 1999.

[AbS] S. S. Abhyankar and G. S. Sundaram, Galois theory of Moore–Carlitz–Drinfeld modules, *C. R. Acad. Sci. Paris* **325** (1997), 349–353.

[Art] E. Artin, The orders of linear groups, *Comm. Pure Appl. Math.* **8** (1955), 355–365.

[BVa] G. D. Birkhoff and H. S. Vandiver, On the integral divisors of $a^n - b^n$, *Ann. of Math.* **5** (1904), 173–180.

[Cam] P. J. Cameron, Finite permutation groups and finite simple groups, *Bull. London Math. Soc.* **13** (1981) 1–22.

[Car] L. Carlitz, A class of polynomials, *Trans. Amer. Math. Soc.* **43** (1938), 167–182.

[Dem] P. Dembowski, "Finite Geometries," Springer-Verlag, Berlin/NewYork, 1968.

[Dic] L. E. Dickson, On the cyclotomic function, *Amer. Math. Monthly* **12** (1905), 86–89.

[Fei] W. Feit, On large Zsigmondy primes, *Proc. Amer. Math. Soc.* **102** (1988), 29–36.

[Gos] D. Goss, "Basic Structures of Function Field Arithmetic," Springer-Verlag, Berlin/ New York, 1996.

[Hay] D. R. Hayes, Explicit class field theory for rational function fields, *Trans. Amer. Math. Soc.* **189** (1974), 77–91.

[He1] C. Hering, Transitive linear groups and linear groups which contain irreducible subgroups of prime order, *Geom. Ded.* **2** (1974), 425–460.

[He2] C. Hering, Transitive linear groups and linear groups which contain irreducible subgroups of prime order II, *J. Algebra* **93** (1985), 151–164.

[Hup] B. Huppert, "Endliche Gruppen I," Springer-Verlag, Berlin/New York, 1983.

[Kan] W. M. Kantor, Homogeneous designs and geometric lattices, *J. Combin. Theory Ser. A* **38** (1985), 66–74.

[Lie] M. W. Liebeck, The affine permutation groups of rank three, *Proc. London Math. Soc.* **54** (1987), 477–516.

[LNi] R. Lidl and H. Niederreiter, "Finite Fields," Addison–Wesley, ReedMS, MA, 1983.

[Moo] E. H. Moore, A two-fold generalization of Fermat's theorem, *Bull. Amer. Math. Soc.* **2** (1896), 189–199.

[Sin] J. Singer, A theorem in finite projective geometry and some applications in number theory, *Trans. Amer. Math. Soc.* **43** (1938), 377–385.

[Zsi] K. Zsigmondy, Zur Theorie der Potenzreste, *Monatsch. Math. Phys.* **3** (1892), 265–284.